

团 体 标 准

T/CAAMM xxxx—20xx

智能农机装备数字孪生系统

第 3 部分：评价方法

Digital twin system of intelligent agricultural machinery

part 3: Evaluation method

（报批公示稿）

202x-xx-xx 发布

202x-xx-xx 实施

中国农业机械工业协会 发 布

目 次

前言II

1 范围.....1

2 规范性引用文件.....1

3 术语和定义.....1

4 概述.....2

5 数字孪生系统总体评价方法.....2

6 数字孪生系统构成要素评价方法.....9

前 言

本文件按照 GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》给出的规则起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由中国农业机械工业协会提出。

本文件由中国农业机械工业协会归口。

本文件起草单位：中国农业大学、北京农学院、洛阳智能农业装备研究院有限公司、北京市农林科学院智能装备技术研究中心、中国农业机械化科学研究院集团有限公司、洛阳拖拉机研究所有限公司、北京启维数字科技有限公司。

本文件主要起草人：杜岳峰、郭大方、宋正河、陈度、郭志强、黄胜操、尹彦鑫、周立明、陈凯康、王东青、高辽远、吴传鑫、栗晓宇、温昌凯、武秀恒、乔智、王林泽、吴志康、马若飞。

本文件为首次发布。

智能农机装备数字孪生系统 第3部分:评价方法

1 范围

本文件描述了智能农机装备数字孪生系统（以下简称“数字孪生系统”）的评价方法。
本文件适用于数字孪生系统的规划、开发、管理、运维、评价和验收等工作。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期的对应版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 41723 自动化系统与集成 复杂产品数字孪生体系架构

GB/T 43441.1 信息技术 数字孪生 第1部分：通用要求

3 术语和定义

GB/T 41723 和 GB/T 43441.1 界定的术语和定义适用于本文件。

3.1

建模仿真 modeling and simulation

建模仿真是指将现实中的农机装备及其作业过程，依据其结构、功能和行为特征，通过数学、物理、逻辑等方法建立虚拟模型，并在计算机平台上进行动态仿真运行，以预测其性能、分析其行为、优化其设计或支持其控制决策的技术过程。

3.2

多源异构数据 multi-source heterogeneous data

在同一系统或应用中，来源于不同数据源、采用不同格式、结构和类型的多种数据集。数据具有不同的维度、不同的数据类型和不同的来源。

3.3

数据交互 Multi-source Heterogeneous Data

指数字孪生系统中虚拟模型与物理实体之间，以及系统各组件之间进行数据采集、传输、同步与反馈。

4 概述

在《智能农机装备数字孪生系统 第1部分：参考架构》中，规定了数字孪生系统由物理实体、虚

拟模型、数据与连接、支持服务、业务应用五个核心要素组成，构成系统整体架构，如图 1 所示。

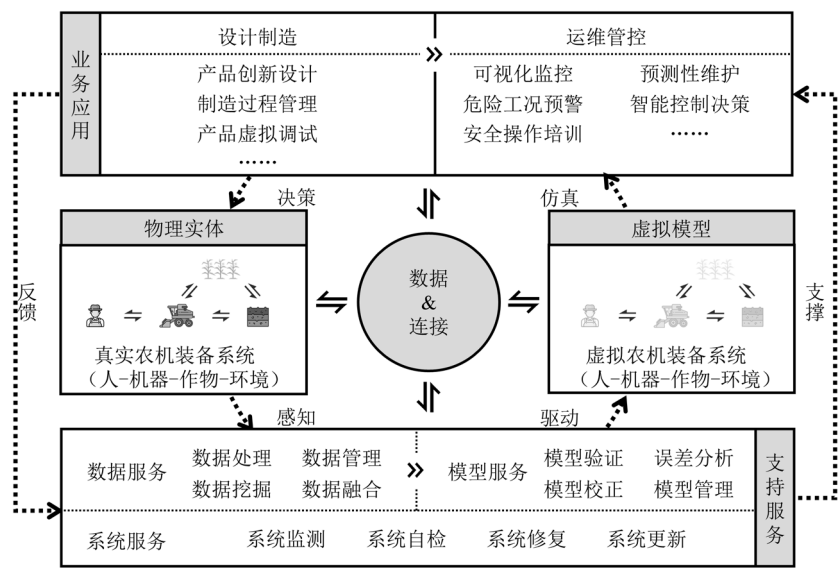


图1 数字孪生系统总体框架

在《智能农机装备数字孪生系统 第 2 部分：技术要求》中，规定了系统的技术要求涵盖功能、性能、用户体验、安全、运维五个维度，构成系统技术指标评价体系，如图 2 所示。

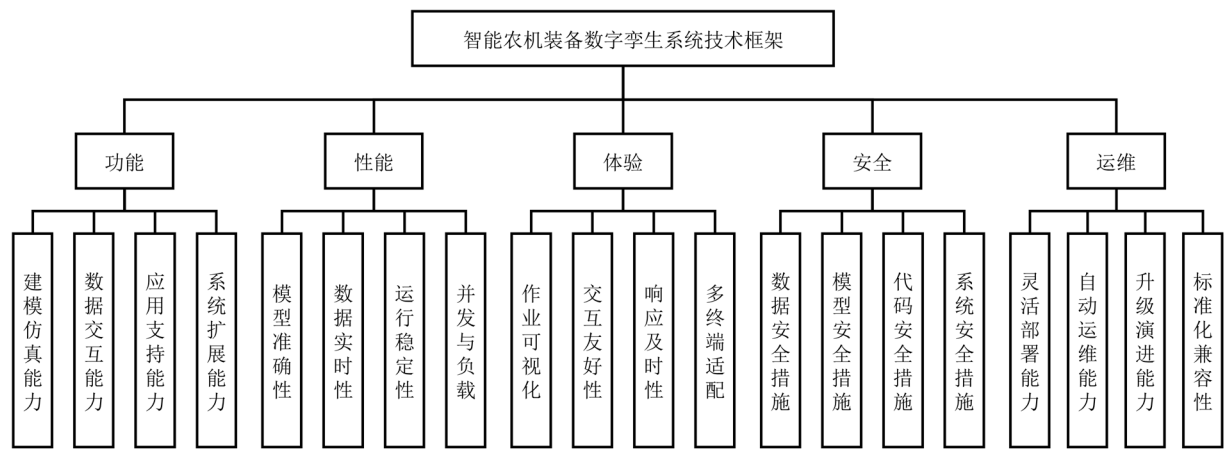


图2 系统技术指标评价维度

本文件在前置文件的基础上，提出了相应的评价方法体系。标准内容围绕上述五个评价维度，分别从以下两个层面进行规定：

- a) 系统层面：面向整个数字孪生系统，提出其在功能、性能、用户体验、安全、运维五个维度的通用评价方法；
- b) 要素层面：面向构成系统的五个核心要素——物理实体、虚拟模型、数据与连接、支持服务、业务应用，分别提出其在上述五个维度的具体评价方法。

5 数字孪生系统总体评价方法

5.1 系统功能评价方法

- a) 建模仿真能力评价

- 方法：基于虚拟模型开展仿真分析，通过专家评审与关键指标分析，考察模型在结构表达和动态响应方面的合理性与可信度。
- 所用设备：仿真平台软件、农机实测数据采集系统、误差评价工具、专家评审表。
- 步骤：
 - 选择典型作业场景（如播种、耕整、植保、收获等）；
 - 在系统中加载仿真模型，覆盖动力系统、传动系统、作业机构等关键部件；
 - 运行虚拟仿真并采集关键指标输出（如速度、轨迹、姿态等）；
 - 同步采集实地作业数据，进行数据对比分析；
 - 计算 RMSE、误差百分比、相关系数、动态响应时间差等指标，辅助评价响应特性；
 - 对模型结构、行为逻辑与响应特征进行定性评审；
 - 综合以上各项指标的综合得分（可采用加权平均法）得到建模仿真能力的最终评价结果。
- b) 数据交互能力评价
 - 方法：验证系统是否支持虚实之间的实时/周期性数据通信、状态同步及反馈控制，确保虚拟系统反映真实作业状态。
 - 所用设备：网络分析工具、抓包软件（如 Wireshark）、CAN 总线分析仪、边缘计算设备。
 - 步骤：
 - 配置农机设备与数字孪生系统的通信接口；
 - 执行指定作业任务并记录数据交互日志；
 - 使用工具抓取通信流，检测数据传输的延迟、丢包与同步更新频率；
 - 验证模型状态随实物变化的同步准确性，计算状态同步误差，即虚拟与真实关键变量（位置、姿态、速度）的均方根误差；
 - 多次下发控制指令，观察虚实系统反馈一致性，计算控制反馈一致率，即控制指令下发后虚实执行结果一致的比例。
 - 综合以上各项指标的综合得分（可采用加权平均法）得到数据交互能力的最终评价结果。
- c) 应用支持能力评价
 - 方法：检验系统是否提供覆盖设计、制造、运维、培训等阶段的典型应用功能，功能完备性和适用性。
 - 所用设备：业务流程测试用例、系统用户操作日志分析工具、用户使用满意度调查问卷。
 - 步骤：
 - 列出平台所宣称支持的业务应用场景（如预测性维护、虚拟调试等）；
 - 进入各模块逐一检验功能可用性与流程完整性；
 - 执行关键操作路径并记录响应时效与功能完整度；
 - 开展用户调研，收集典型用户群体的满意度反馈；
 - 统计覆盖率和满意度得分，形成综合评价结论。
- d) 系统扩展能力评价
 - 方法：评价系统的模块化结构、平台兼容性与 API 开放性，是否支持外部系统集成与后续功能拓展。
 - 所用设备：API 调用工具（如 Postman）、第三方系统模拟平台、接口协议一致性测试软件。
 - 步骤：
 - 查阅系统架构文档，确认是否具备模块化与分层设计；

- 测试 API 是否符合 RESTful 或 OpenAPI 规范，验证其功能描述完整性；
- 模拟第三方系统调用接口，实现数据读写联通；
- 统计接口响应时间、失败率与兼容平台数量；
- 评价接口使用文档的完整性与开发支持服务情况。

5.2 系统性能评价方法

a) 模型准确性评价

- 方法：通过虚实对照实验，通过关键参数对比与误差指标计算，评价虚拟模型对实际农机装备运行状态的反映能力与精度，确保模型误差控制在可接受范围内。
- 所用设备：农机实测数据采集装置（如 GNSS、IMU、传感器等）、仿真平台、误差分析软件（如 MATLAB、Python 脚本等）。
- 步骤：
 - 在标准试验田中部署农机装备，安装用于轨迹、速度、姿态等参数采集的传感器；
 - 在系统中调用对应虚拟模型，并加载相同工况条件进行仿真；
 - 获取实际与仿真结果的关键参数数据（如路径、速度、扭矩、作业幅宽等）；
 - 使用误差评价工具计算 RMSE、最大绝对误差、误差分布等指标；
 - 判断模型误差是否在业务定义的阈值范围内（如 $<5\%$ ）；
 - 输出模型在特定工况下的准确性评价结论。

b) 数据实时性评价

- 方法：测量系统在采集、传输、处理及反馈数据过程中的整体响应时间，评价其在动态环境中的时效适应性。
- 所用设备：数据流生成器、网络抓包工具（如 Wireshark）、延迟测量脚本、边缘节点设备。
- 步骤：
 - 在农机作业过程中采集多源异构数据（如作业状态、环境参数等）。
 - 使用抓包工具记录从采集端→边缘处理→平台入库全过程时间戳。
 - 对每一次数据更新计算传输延迟与系统响应时间。
 - 重复测试不少于 100 组，统计平均时延、最大时延、抖动等指标。
 - 对比业务场景的响应时效要求（如 $\leq 1s$ ）判断达标情况。

c) 运行稳定性评价

- 方法：验证系统在高强度、长时间运行条件下的稳定性与容错能力，防止出现崩溃、卡顿或功能中断。
- 所用设备：模拟器、多线程负载生成器、长时间作业测试工具、故障注入平台。
- 步骤：
 - 在典型作业负载下运行系统不少于 8 小时，持续监测 CPU、内存、网络与磁盘使用率。
 - 检查系统是否发生异常终止、内存泄漏或响应中断。
 - 在运行过程中插入异常事件（如网络中断、数据包丢失等）测试系统容错恢复能力。
 - 检查是否支持断点续传、边缘缓存机制等稳定性设计。
 - 汇总测试日志，评价系统稳定性指标，如 MTBF（平均无故障时间）等。

d) 并发与负载能力评价

- 方法：评价系统在多设备同时接入、多任务并发运行条件下的处理能力，验证其支持协同作

业的能力。

- 所用设备：模拟多台农机的虚拟接入平台、并发请求生成器、服务器性能监控工具。
- 步骤：
 - 同时模拟不少于 5 台智能农机设备的数据接入，包括位置信息、作业状态与环境数据。
 - 发送并发控制指令与数据请求，观察系统是否能稳定响应。
 - 使用监控工具记录 CPU、内存、网络带宽等资源占用情况。
 - 逐步增加接入设备数量至系统极限，记录最大并发接入数量与系统崩溃临界点。
 - 判断是否符合“多台农机协同作业”的应用需求（建议并发 ≥ 10 台农机为合格）。

5.3 系统体验评价方法

a) 作业可视化能力评价

- 方法：评价系统是否具备直观的三维作业可视化、轨迹回放、状态监控和故障仿真功能，以提升用户对作业过程的感知与理解。
- 所用设备：三维图形渲染终端、轨迹回放播放器、故障模拟模块、作业状态数据采集系统。
- 步骤：
 - 加载典型作业数据（如播种轨迹、喷雾路径等）到系统中；
 - 调用三维模型界面，观察作业轨迹、机具运动状态是否可视化呈现；
 - 模拟特定故障（如堵种、喷雾异常等），验证是否可动态显示并触发故障响应动画；
 - 检查是否支持作业轨迹回放、状态联动展示（如颜色、箭头、标签变化等）；
 - 组织用户评分视觉清晰度、内容丰富度与操作直观性（建议满分 100 分前提下，得分 ≥ 80 分为合格）。

b) 交互友好性评价

- 方法：评价系统界面设计、语言适配性与操作流程是否简洁，是否适用于农业一线用户群体的使用习惯与文化背景。
- 所用设备：图形界面操作终端、交互操作日志分析工具、多语言设置界面。
- 步骤：
 - 观察系统是否提供图形化界面、图标与文字指引一致、按钮布局合理；
 - 切换语言界面（如简体中文/英文），检查界面信息是否完整翻译、语义通顺；
 - 执行典型作业流程操作（如设备上电 \rightarrow 启动 \rightarrow 查看作业信息），统计平均点击次数与操作时间；
 - 邀请不少于 5 名基层用户进行操作试用，收集其对界面友好性、操作习惯匹配度的反馈；
 - 汇总用户满意度评分，建议满分 100 分前提下，得分高于 80 分视为符合“交互友好性”要求。

c) 响应及时性评价

- 方法：测量用户操作指令到系统界面反馈之间的响应时间与图表刷新周期，确保界面流畅不延迟。
- 所用设备：高频点击模拟工具、系统性能监测器、屏幕录像分析软件。
- 步骤：
 - 在典型操作场景中（如点击“作业详情”按钮），记录从用户点击到界面响应的时间（ms 级）；

- 在运行状态下观察系统图表、数据面板刷新频率，是否小于 2s/次；
- 连续执行 50 次以上操作命令，统计最大、最小与平均响应时间；
- 检查在高负载或弱网络条件下响应是否仍稳定流畅；
- 确认是否达到业务需求中设定的响应时间标准。

d) 多终端适配能力评价

- 方法：验证系统是否支持 PC、移动端、机载终端等多种终端形式访问，确保可在田间与办公室环境灵活部署。
- 所用设备：个人电脑、手机、车载触控终端、不同分辨率显示器。
- 步骤：
 - 在各类型终端中分别部署系统客户端或访问 Web 端；
 - 检查界面是否根据终端分辨率自适应，功能是否完整可用；
 - 执行统一操作流程（如查看作业状态、下发任务等），比较功能是否一致；
 - 记录界面显示异常、操作受限、排版错乱等适配问题；
 - 至少支持三种终端设备，且每种设备完成测试流程无明显功能缺失则评价合格。

5.4 系统安全评价方法

a) 数据安全措施评价

- 方法：评价系统在数据采集、传输与存储全过程中是否采取有效加密手段，确保数据机密性和完整性不被泄露或篡改。
- 所用设备：网络抓包工具（如 Wireshark）、加密传输协议测试平台、安全渗透测试工具。
- 步骤：
 - 检查系统是否启用加密传输协议（如 HTTPS、TLS），并验证密钥长度、加密算法类型；
 - 使用抓包工具对传输数据进行截取，验证是否为加密格式，是否存在泄露风险；
 - 检查数据库或数据缓存中的静态存储数据是否加密存储；
 - 模拟数据窃听与篡改，评价系统是否具备防护机制与报警能力；
 - 检查是否设置数据访问控制、非法请求过滤等基础防护措施。

b) 模型安全措施评价

- 方法：验证系统对虚拟模型在构建、调用、部署、更新等全生命周期中的保护机制，确保模型不被篡改、泄露或替换。
- 所用设备：模型版本控制系统、完整性校验工具（如 MD5/SHA256 等）、访问权限管理系统。
- 步骤：
 - 检查系统是否对模型设有版本控制与变更记录机制；
 - 对关键模型文件进行完整性校验，模拟篡改后验证系统是否能识别异常；
 - 检查模型调用过程是否受限于权限控制（如只读、不可导出、限角色访问等）；
 - 模拟非法替换模型行为，系统是否能拦截、报警或恢复；
 - 审查模型文件的访问日志是否完整、可追踪。

c) 代码安全措施评价

- 方法：对系统代码进行安全审查和漏洞扫描，评价其是否存在已知漏洞、恶意代码或安全风险。
- 所用设备：静态代码分析工具（如 Fortify、SonarQube 等）、漏洞扫描器（如 Nessus 等）、

渗透测试脚本。

— 步骤：

- 使用静态分析工具扫描代码库，识别潜在漏洞（如 SQL 注入、XSS 等）；
- 使用动态漏洞扫描工具测试运行环境中可能存在的系统漏洞；
- 模拟外部攻击（如非法请求注入、目录遍历、命令执行等），评价系统防护反应；
- 检查是否建立安全补丁更新机制，记录历史补丁日志；
- 审核是否存在后门程序、明文密码等高危问题，并形成安全整改清单。

d) 访问权限控制评价

— 方法：评价系统是否实现多级权限控制、角色划分和用户认证机制，有效防止非法访问和误操作。

— 所用设备：用户权限管理模块、认证机制配置界面、多用户角色测试账号。

— 步骤：

- 检查系统用户身份验证机制是否支持多因子认证（如密码+验证码）；
- 查看角色与权限配置界面，是否支持按模块、操作级别划分；
- 使用不同权限账户登录系统，验证其访问范围是否正确隔离；
- 测试是否存在权限越界、任意调用或误操作风险；
- 检查权限变更日志与用户认证日志的完整性和可追溯性。

e) 系统防护能力评价

— 方法：验证系统是否具备防火墙、入侵检测、防病毒、非法访问阻断等防护机制，应对外部攻击与内部泄密风险。

— 所用设备：入侵检测系统（IDS）、防火墙策略模拟器、病毒扫描工具、DoS 攻击模拟平台。

— 步骤：

- 审查系统部署的网络防护设施，包括防火墙、入侵检测和安全策略设置；
- 模拟外部攻击行为（如端口扫描、爆破登录、DDoS 攻击等），测试系统是否能检测与响应；
- 使用病毒扫描工具检测是否存在恶意程序或脚本注入风险；
- 查看系统在受到攻击时是否可自动限制 IP、发送告警、保护数据；
- 审核系统更新周期、安全策略维护周期，确保防护机制持续有效。

f) 审计追溯机制评价

— 方法：检查系统是否具备完整的日志记录与审计追溯机制，支持安全事件的事后分析与责任追踪。

— 所用设备：日志管理系统、安全审计平台、时间戳校验工具。

— 步骤：

- 检查系统是否记录用户登录、操作、数据变更、异常事件等信息；
- 验证日志内容是否包含时间戳、用户 ID、操作对象与动作描述等关键字段；
- 模拟异常操作（如非法导出、强制中断等），查看日志是否完整记录；
- 检查是否支持日志备份、导出、加密存储；
- 使用审计工具进行日志追踪测试，验证其在事件定位与行为重建中的支持能力。

5.5 系统运维评价方法

a) 灵活部署能力评价

- 方法：验证系统是否支持在云端、本地及混合部署环境中灵活部署，适应不同网络环境和应用场景的快速上线需求。
- 所用设备：云服务平台、本地服务器、边缘计算终端、网络环境模拟器。
- 步骤：
 - 在不同环境中分别部署系统组件，包括本地部署、云端部署与边缘混合架构；
 - 检查安装配置流程是否标准化、可自动化（如使用 Docker、K8s 等）；
 - 测试部署完成后的系统访问延迟、运行稳定性和关键功能可用性；
 - 在弱网络或无公网环境下，验证系统的启动能力与基础功能可运行性；
 - 评价部署文档的完备性与部署时间的控制情况（如≤2 小时完成初始部署）。

b) 自动运维能力评价

- 方法：评价系统是否具备自动监控、故障检测、资源调度、日志管理等自动化运维能力，降低人工维护负担。
- 所用设备：系统监控平台（如 Prometheus+Grafana 等）、运维告警系统、故障注入工具、日志采集系统。
- 步骤：
 - 验证系统是否自动采集 CPU、内存、存储、带宽等运行指标；
 - 模拟 CPU 过高、网络异常等典型故障场景，检查系统是否能自动识别并发出报警；
 - 检查是否支持邮件、短信或 APP 通知等多渠道报警机制；
 - 检查是否支持自动重启、资源清理、负载均衡等应对机制；
 - 审核日志采集策略和周期，是否支持自动归档、压缩、分级管理。

c) 升级演进能力评价

- 方法：验证系统是否支持模块化架构、热更新以及平滑升级等机制，确保版本演进不中断业务运行。
- 所用设备：版本管理平台、灰度发布工具（如 K8s 滚动升级、蓝绿部署系统等）、业务模拟终端。
- 步骤：
 - 检查系统是否按模块组织部署包，支持单模块独立升级；
 - 模拟版本迭代过程（如 v1.0→v1.1），验证是否可在不中断服务的前提下完成更新；
 - 检查更新过程是否影响用户连接、系统响应或作业执行；
 - 检查是否支持升级前版本回退机制及升级失败处理逻辑；
 - 审核升级日志与变更文档是否齐全，版本记录是否清晰。

d) 标准化兼容性评价

- 方法：评价系统是否符合相关农业物联网、农机信息化等行业标准，具备良好的跨平台运行和设备兼容能力。
- 所用设备：通信协议分析仪、农机信息编码测试平台、行业标准接口测试工具。
- 步骤：
 - 检查系统是否支持主流农业物联网通信协议（如 Socket、MQTT、CoAP 等）与标准数据格式（如 JSON、XML 等）；
 - 验证系统接口是否符合农机编码规范；

- 测试与典型农机设备、农机管理平台的数据交换接口，记录通信成功率与接口适配率；
- 检查系统是否支持多种操作系统、多种数据库与中间件兼容性；
- 提供符合标准的 API 文档（如 OpenAPI 3.0 等），便于外部集成与系统对接。

6 数字孪生系统构成要素评价方法

6.1 物理实体评价方法

a) 功能评价方法

- 方法：评价物理实体的多维度感知能力、传感器系统配置情况及执行单元的响应与控制能力，确保其满足数字孪生系统数据来源和动作执行基础。
- 所用设备：多通道数据采集器、工业级传感器验证平台、信号测试仪、执行动作记录装置（含响应计时模块）。
- 步骤：
 - 配置传感器采集系统，记录位置、速度、姿态等关键运行参数；
 - 检查是否具备关键传感信息融合机制，通过故障注入方式验证感知连续性；
 - 对目标设备执行软测量模型验证，测试对不可测物理量（如土壤紧实度等）的估算精度；
 - 检查传感器规格参数、认证标识，验证其抗震动、耐高湿、防尘性能；
 - 操作执行单元，验证是否具备闭环控制、远程控制（线控底盘）和响应时间满足任务需求；
 - 模拟多任务调度场景，测试是否具备执行动作优先级控制和冲突避免机制。

b) 性能评价方法

- 方法：测试物理实体的数据采集频率、响应延迟、控制准确性及环境适应能力，确保其在田间作业下具备高性能。
- 所用设备：高速数据记录仪、控制指令发射器与响应检测计、硬件试验平台、标准协议测试仪。
- 步骤：
 - 设置多种作业模式（高速行驶/低速作业），检测各工况下采集频率是否自适应调整；
 - 测试参数个性化设置功能，记录不同传感器频率设定的精度与稳定性；
 - 验证时间戳校准机制，确保多通道数据在同一参考时钟下同步；
 - 使用响应检测设备测试控制延迟是否满足<100ms（关键控制<50ms）要求；
 - 模拟电磁干扰场景，测试抗干扰能力与丢包率（建议<0.01%）；
 - 在高温、高湿、震动测试环境下运行设备，记录性能变化，判断是否符合 GB/T 2423 系列相关等级。

c) 体验评价方法

- 方法：通过对执行状态可视化、告警提示、交互反馈机制等的评价，判断物理实体对操作者的信息传递是否及时、直观。
- 所用设备：操作终端界面、语音/灯光/图标提示装置、用户观察记录表、现场测试用农机实物。
- 步骤：
 - 模拟感知异常（如信号丢失、误差漂移等），检查系统是否触发视觉/听觉/图标提示；
 - 记录信息传递时间、反馈方式及覆盖参数是否全面；

- 邀请典型用户（如驾驶员、维保人员等）现场体验并打分，收集反馈意见；
- 汇总反馈结果并进行体验满意度量化评分（建议满分 100 分的前提下，评分 ≥ 80 分为合格）。

d) 安全评价方法

- 方法：从感知安全、执行安全、虚实同步异常防护、指令隔离与通信安全等多个方面对物理实体安全能力进行全面测试。
- 所用设备：异常传感器信号注入器、安全控制策略管理系统、通信加密模块测试平台、同步误差测距装置、控制器日志分析工具。
- 步骤：
 - 注入异常传感器数据（如掉线、漂移、超限等），检测是否自动识别异常状态并告警或触发降级控制；
 - 模拟控制指令越界、执行失败等情况，确认是否具备指令校验机制与反馈闭环；
 - 设置关键控制动作（如紧急停车）误发场景，观察设备是否阻断并警告；
 - 在虚拟模型中手动调节位置，检测物理设备虚实误差报警触发门限（如 10cm 内是否预警）；
 - 尝试远程发送控制指令，检查是否具备操作双重确认机制、指令白名单与权限认证流程；
 - 检查通信加密方式，记录是否支持 TLS、数字签名、数据校验（如 CRC 等）；
 - 查看所有远程控制指令、异常响应记录是否具备完整日志和追溯性。

e) 运维评价方法

- 方法：验证物理实体设备的模块化维护、状态监控、兼容适配与远程诊断能力，确保其具备高可维护性和适应性。
- 所用设备：模块热插拔测试平台、远程维护工具包、诊断支持平台、边缘计算节点、设备兼容性测试装置。
- 步骤：
 - 拆卸传感器与执行器部件，验证是否可快速更换、即插即用（无需重配置）；
 - 检查控制器是否支持本地按钮调试与远程参数修改；
 - 启动设备，查看是否可实时远程监测运行状态、上报故障码、发起寿命评价；
 - 模拟设备故障，记录预警是否及时准确，是否支持本地+远程联合诊断功能；
 - 加载不同品牌农机设备通信协议，测试兼容性与多协议解析能力；
 - 插入新增模块（如新型传感器），测试系统是否自动识别并加载默认配置；
 - 验证历史日志可否在边缘侧缓存并上传平台，实现故障后回溯与分析支持。

6.2 虚拟模型评价方法

a) 功能评价方法

- 方法：验证虚拟模型是否支持多物理场、多尺度建模，具备良好的交互性、场景适应能力、模块化设计和数据驱动能力，满足动态仿真和业务应用需求。
- 所用设备：联合仿真平台（如 Simulink、Modelica、AMESim、ANSYS 等）、图形化建模接口、参数自动匹配规则库、版本管理系统、机器学习模块。
- 步骤：
 - 加载涵盖机械、电气、液压、土壤-作物-装备耦合等子模型，测试其是否可在多物理域

中协同运行；

- 检查建模是否支持从部件级（如单一传感器）到系统级（整车系统）的分层结构；
- 测试是否支持 FMU 标准接口，能否与多平台进行模型互操作与联合仿真；
- 在仿真 GUI 中运行模型，验证是否支持用户参数加载、配置、动态调参与运行控制；
- 输入典型作业任务（播种、收获等），检测模型是否根据任务自适应切换结构或参数；
- 对模型进行组件级封装，检查是否支持模板化调用、版本追踪、依赖检测与耦合仿真；
- 加载历史作业数据，测试模型是否具备数据驱动的结构/参数更新能力，并查看是否记录更新过程、训练样本及模型演化日志。

b) 性能评价方法

- 方法：测试虚拟模型的仿真精度、响应延迟、资源占用与部署能力，确保其运行稳定、效率高、适配性强。
- 所用设备：高精度仿真结果对比系统、实时性测试框架（如 dSPACE、LabVIEW 等）、轻量化建模工具（如代理模型平台）、边缘计算网关与版本同步模块。
- 步骤：
 - 在典型作业任务中运行仿真，记录关键性能指标（如路径跟踪误差、执行偏差等）与实际值对比，评价仿真精度是否在误差阈值内；
 - 测试仿真模型在不同复杂度下的响应时延与循环频率，确保其满足业务实时性需求（如闭环控制<100ms）；
 - 检查是否具备状态缓存、异步计算能力，在大模型负载下依然平稳运行；
 - 切换轻量版/标准版/高精度版模型，测试资源消耗变化，记录各配置下内存占用与 CPU 利用率；
 - 验证模块按需加载、资源动态调度等功能，统计模型调度效率；
 - 将轻量模型部署至边缘节点，进行仿真或诊断任务，测试断点续传、差异同步、版本回滚等功能；
 - 检查是否具备模型同步一致性校验功能，如 hash 对比、双向更新记录。

c) 用户体验评价方法

- 方法：评价模型在交互操作、加载配置、结构可视化、参数输入等方面的用户友好性与灵活性。
- 所用设备：仿真界面终端、参数模板管理工具、用户操作日志记录系统。
- 步骤：
 - 运行模型界面，观察是否支持可视化结果展示、结构组件启停、参数动态调整等功能；
 - 记录操作路径（加载→配置→运行→结果查看）所需步骤数与平均耗时；
 - 检查是否具备参数自动填写/导入模板、模型切换引导界面等功能；
 - 评价是否支持结构与数据的可视对比、关键参数高亮显示等增强用户体验的设计；
 - 组织用户试用并打分，收集对操作流程、交互逻辑的满意度评价。

d) 安全评价方法

- 方法：检查模型在调用、数据存储、接口访问方面的安全控制措施，包括日志记录、数据加密与权限控制等。
- 所用设备：模型日志审计系统、安全协议测试平台、接口认证工具（如 OAuth2.0 验证器）、加密算法测试模块。

— 步骤：

- 调用模型服务多次，验证是否自动生成日志记录，包括调用时间、用户、参数、结果状态等内容；
- 检查日志记录是否支持按字段、用户、时间范围进行查询和导出，是否能与告警系统联动；
- 抓包检查模型结构文件与数据在传输过程中是否采用 TLS 1.2 以上协议加密；
- 检查模型核心参数是否存储为 AES-256 加密格式，是否具备访问权限控制机制（如只读、限角色访问）；
- 验证模型结构与配置文件是否支持完整性校验（如 hash 签名）和异常变更检测；
- 测试接口访问控制是否启用 OAuth2.0 认证，是否具备访问频率限制策略；
- 模拟高风险操作（如强制模型重置），检查是否启用管理员授权或双因子验证流程。

e) 运维评价方法

— 方法：评价模型的版本管理、运行状态监控、自诊断能力及平台兼容性，确保其可维护、可追踪、可扩展。

— 所用设备：模型版本管理平台、运行状态监控控制台、自诊断工具、兼容性测试系统、主流仿真平台（如 Simulink、Modelica 等）。

— 步骤：

- 检查是否支持模型的创建、发布、修改、更新记录版本信息，支持状态标识与回滚；
- 测试模型操作日志记录功能是否完整，记录包括操作人、时间、内容等；
- 打开模型运行状态监控界面，查看是否显示响应时间、执行频率、资源使用等指标；
- 模拟模型运行异常（如计算中断），系统是否自动诊断原因、发出告警、切换备用模型；
- 在不同仿真平台上部署模型，验证其在 FMU 或联合仿真环境下的兼容性；
- 同时运行多个异构模型，测试混合运行与并发能力，记录运行稳定性与资源调度效率；
- 验证平台是否支持新增仿真模块的快速集成与横向扩展机制。

6.3 数据与连接评价方法

a) 功能评价方法

— 方法：评价系统对多种通信协议的支持程度、“农机-模型”之间的联动能力，以及外部数据源的接入与处理能力。

— 所用设备：协议分析仪、Modbus/CAN 模拟器、API 调试工具（如 Postman 等）、模拟外部数据平台（如气象服务）、数据预处理模块。

— 步骤：

- 测试通信模块是否支持主流协议（如 Modbus RTU/TCP、CAN、Ethernet 等），验证物理层连接及报文识别；
- 插拔通信设备，检查系统是否能自动识别并重新初始化驱动，验证热插拔支持；
- 查看通信驱动模块结构，测试是否支持插件式加载与协议扩展；
- 模拟作业任务，在农机装备运行过程中观察模型状态同步情况；
- 设置虚拟模型输出（如路径规划），验证指令是否能成功转化为设备控制指令并下发执行；
- 配置不同联动规则（如“深度>20cm 则触发模式切换”等），测试阈值-反馈-执行的闭环

环运行情况；

- 接入气象/土壤第三方数据源，通过 RESTful API 发送数据，验证系统是否能接收、解析、标准化并应用；
- 模拟数据缺失场景，检查是否自动补全、同步或修复异常时间戳。

b) 性能评价方法

- 方法：评价数据采集、双向传输、边缘预处理等能力，确保数据链路在实际作业中高效、稳定运行。
- 所用设备：高频数据采集终端、通信负载模拟器、边缘计算单元、延迟测试器、抓包工具。
- 步骤：
 - 设置不同作业模式（低速/高速/间断），验证采集频率是否自适应调整；
 - 同时测试指令下发与数据上传通道，记录双向传输时延，目标响应<200ms；
 - 验证所有采集数据是否附带时间戳，是否使用标准结构（如 JSON/XML/Binary）；
 - 模拟网络不稳定或高负载情况，观察是否优先保障关键数据（如位置、速度等）的刷新频率；
 - 在边缘设备部署轻量 AI 模型，实时推理并生成诊断标签，测试其准确性与时效性；
 - 验证系统能否在带宽受限时，自动切换数据处理重心到边缘端，保持核心服务不中断。

c) 用户体验评价方法

- 方法：从配置便捷性、连接状态展示、交互性和多终端一致性等角度，评价数据通信的可视性与用户感知体验。
- 所用设备：图形化配置界面、移动端与 PC 终端、通信状态监控平台、数据展示交互系统。
- 步骤：
 - 在配置界面中测试通信频率、字段、采集策略的导入/导出功能，检查是否支持模板化调用；
 - 检查是否支持按设备类型推荐默认配置模板，并对配置过程进行可视引导；
 - 在系统状态栏中查看设备在线状态、信号强度、延迟曲线，验证展示的实时性与直观性；
 - 故意制造断链/丢包/延迟，观察系统是否能以图标、颜色、图表形式准确预警；
 - 浏览实时数据、历史记录，并测试字段筛选、时间轴调整等交互功能；
 - 切换表格、图表、地图等多种展示视图，确认功能一致性与视图适配性；
 - 在 PC 端、移动端、车载终端操作相同功能（如查看通信状态、修改采集参数），验证操作逻辑与界面布局一致性；
 - 模拟弱网环境，验证移动端是否自动缓存数据并在网络恢复后自动补传。

d) 安全评价方法

- 方法：从网络安全、数据加密与完整性、断点续传等方面，评价数据通信链路的安全防护能力与异常容错性。
- 所用设备：防火墙策略配置平台、入侵检测系统、TLS/加密测试工具、CRC 校验模块、链路断开模拟器。
- 步骤：
 - 检查是否部署 IP 白名单、端口过滤、防协议滥用策略，测试策略生效性；
 - 使用 IDS 工具模拟 DDoS 攻击、指令篡改等行为，观察是否能自动识别并联动告警模块；
 - 抓取通信流量，确认是否启用 TLS 1.2 或以上协议，数据内容加密方式为 AES-256；

- 检查所有数据包是否附带完整性校验信息，模拟数据篡改后验证是否能检测与拦截；
- 模拟断网场景，验证本地缓存是否启动，缓存上限是否 ≥ 24 小时；
- 重连后测试 ACK 确认机制与自动重传机制，确认数据完整性；
- 验证断点续传机制是否能准确恢复模型同步状态，避免丢帧或状态漂移。

e) 运维评价方法

- 方法：评价系统在通信链路监控、协议驱动扩展、接口兼容、备份与恢复等方面的运行维护能力。
- 所用设备：通信状态监测系统、驱动热更新框架、兼容性测试平台、备份/恢复工具包、日志审计终端。
- 步骤：
 - 在运维平台查看通信状态图表，检查是否包含连接状态、延迟、丢包率、告警等指标；
 - 人为制造通信异常（如接收阻塞），验证系统是否能实时推送告警；
 - 新增设备类型或协议（如 LoRa），测试是否可热加载驱动插件、无需重启；
 - 调用外部系统 API 接口（如农业数据平台），测试其兼容性与规范性；
 - 检查系统是否支持按设备/时间段/数据类型等维度进行数据恢复；
 - 模拟边缘节点/云端数据丢失场景，测试备份数据恢复流程及精度；
 - 检查是否启用异地容灾备份机制，保留周期是否满足 ≥ 180 天；
 - 验证恢复过程是否支持可视化选择、操作日志记录与快速回滚。

6.5 业务应用评价方法

a) 功能评价方法

- 方法：评价业务应用模块是否覆盖核心作业功能，支持多终端接入、开放集成、模块更新等应用级能力，满足实际农业作业与运营需求。
- 所用设备：仿真交互终端、路径规划引擎、预警模型训练平台、作业历史数据管理系统、API 测试工具（如 Postman 等）、多端访问终端（如个人计算机、移动终端、车载终端等）。
- 步骤：
 - 检查是否支持作业状态、轨迹、传感器数据等多源信息融合展示，验证作业可视化模块是否具备视频流嵌入与回放功能；
 - 运行仿真交互模块，检查是否支持虚拟操作与参数调整，观察是否可与实机状态联动；
 - 启动路径规划引擎，输入障碍物/地形信息，评价自动路径生成、避障处理与多机协同策略效果；
 - 加载作业历史数据与规则库，检查是否能基于 AI/规则触发异常预警与故障预测；
 - 测试基于作业历史与环境参数的作业策略推荐功能，包括时间建议、速度配置等；
 - 启用虚拟培训模块，检验是否具备考核评分、虚拟任务反馈与用户分级评价功能；
 - 在 PC、移动、车载端同时登录系统，验证界面、数据、交互是否保持一致；
 - 测试离线场景下移动端是否能本地缓存数据、浏览模型并在联网后自动同步；
 - 加载不同地图服务（如高德、百度、离线地图等），检查地图服务切换兼容性；
 - 使用 Postman 测试 API 接口，验证数据查询、模型调用、状态上报、指令下发等功能是否可用；
 - 检查 SDK 文档覆盖语言（Python/Java/C#），调用示例是否清晰，错误码说明是否完整；

- 配置 Webhook 推送路径，测试是否能将预警或控制指令实时推送至第三方平台；
- 模拟热更新模块注入过程，观察是否能模块级部署更新而不影响其他服务；
- 测试第三方模块通过沙箱运行，确认其不会影响主系统的稳定性。

b) 性能评价方法

- 方法：测试业务应用的加载响应时间、高可用与故障转移能力、并发处理与系统弹性扩展能力。
- 所用设备：性能测试平台（如 JMeter、Locust）、容器管理平台（Docker/Kubernetes）、监控与日志平台（Prometheus、Grafana）。
- 步骤：
 - 清除缓存后首次访问系统，记录页面加载时间，确保 $\leq 3s$ ；
 - 执行标准操作路径（点击、检索、配置），记录用户操作响应时间，应 $\leq 500ms$ ；
 - 对数据查询接口进行分页、模糊与缓存测试，记录接口响应平均耗时是否 $\leq 200ms$ ；
 - 检查是否启用预加载机制，提前加载关键数据模块；
 - 模拟主节点故障，验证系统是否在 15s 内完成主备切换；
 - 配置 K8s 健康探针，手动关闭服务进程，检查是否自动剔除并恢复副本；
 - 模拟接口异常、任务阻塞等场景，检查是否自动触发服务熔断、性能降级机制；
 - 进行高并发压力测试，验证系统是否能弹性调度 Pod 或容器副本；
 - 对耗时任务（如模型训练、历史回放等）使用异步任务队列处理，确保主线程响应不被阻断；
 - 检查所有微服务是否具备容器化部署能力，验证其部署和伸缩效率。

c) 安全评价方法

- 方法：从权限控制、加密通信、日志审计和终端访问等角度，验证业务系统安全保障机制的健全性和有效性。
- 所用设备：权限管理平台、安全认证系统（OAuth 2.0、JWT）、HTTPS 抓包工具、审计日志分析工具、终端访问白名单系统。
- 步骤：
 - 检查权限配置是否基于 RBAC 模型，是否支持角色分级与模块级权限分配；
 - 测试权限配置界面是否可视化管理，是否支持导出备份与快速复制策略；
 - 验证所有通信是否基于 HTTPS（TLS 1.2+），通过抓包确认数据传输加密有效；
 - 测试终端登录流程，检查是否强制身份认证，支持 OAuth 2.0 或 JWT 认证；
 - 对远程控制等高风险操作，验证是否启用双因素认证或二次确认机制；
 - 查看用户操作日志记录是否完整，包含操作人、时间、参数、IP 与结果等；
 - 审计日志保留是否 ≥ 180 天，并支持关键词检索、导出；
 - 模拟高频操作与越权访问，验证是否触发行为异常识别与封锁机制；
 - 设置终端绑定机制，测试非授权设备访问时的拦截与告警流程；
 - 配置终端白名单，验证数据是否仅限授权设备接入；
 - 配置终端访问策略（只读/读写），并测试不同等级终端的访问能力与限制行为。

a) 用户体验评价方法

- 方法：通过多终端一致性、界面交互友好性、用户上手难度与个性化配置功能，综合评价用户体验质量。

- 所用设备：个人计算机端、移动终端、车载终端，响应式界面测试平台，用户问卷系统，定制化 UI 模块。
 - 步骤：
 - 在多个终端访问系统，执行同一业务流程，确认功能一致性与操作流程一致性；
 - 检查界面是否具备响应式设计，适配屏幕分辨率 $\geq 1280 \times 720$ 的不同终端；
 - 模拟弱网环境，观察是否支持断点续传、离线缓存、网络自动切换；
 - 测试拖拽、滑动、点选等界面交互操作是否顺畅，帧率是否 $\geq 30\text{fps}$ ；
 - 检查图标、配色、字体风格是否统一，并在强光环境下测试可视性；
 - 登录新账户，是否自动触发界面引导、悬浮提示、新手操作流程；
 - 检查术语是否通俗易懂，是否具备“新手/专家模式”切换功能；
 - 自定义仪表板布局、快捷入口、功能收藏项，并验证保存与还原功能是否生效；
 - 配置不同场景下的仿真策略与路径模板，检查系统是否自动调用用户设定方案；
 - 测试主题切换（浅/深色）、字体大小、模块布局等 UI 个性化配置项是否生效、保存与加载顺畅。
-